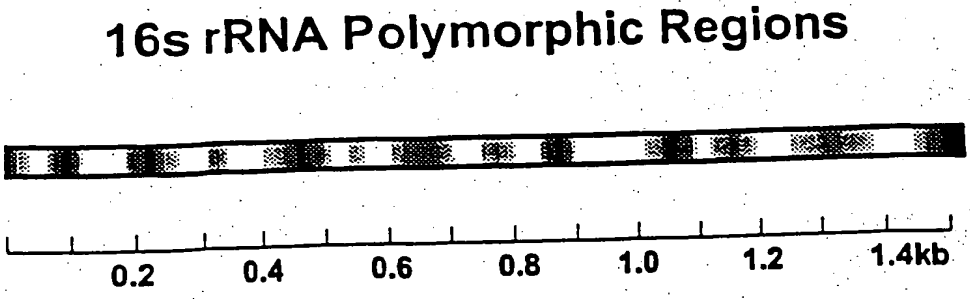


FIG. 1

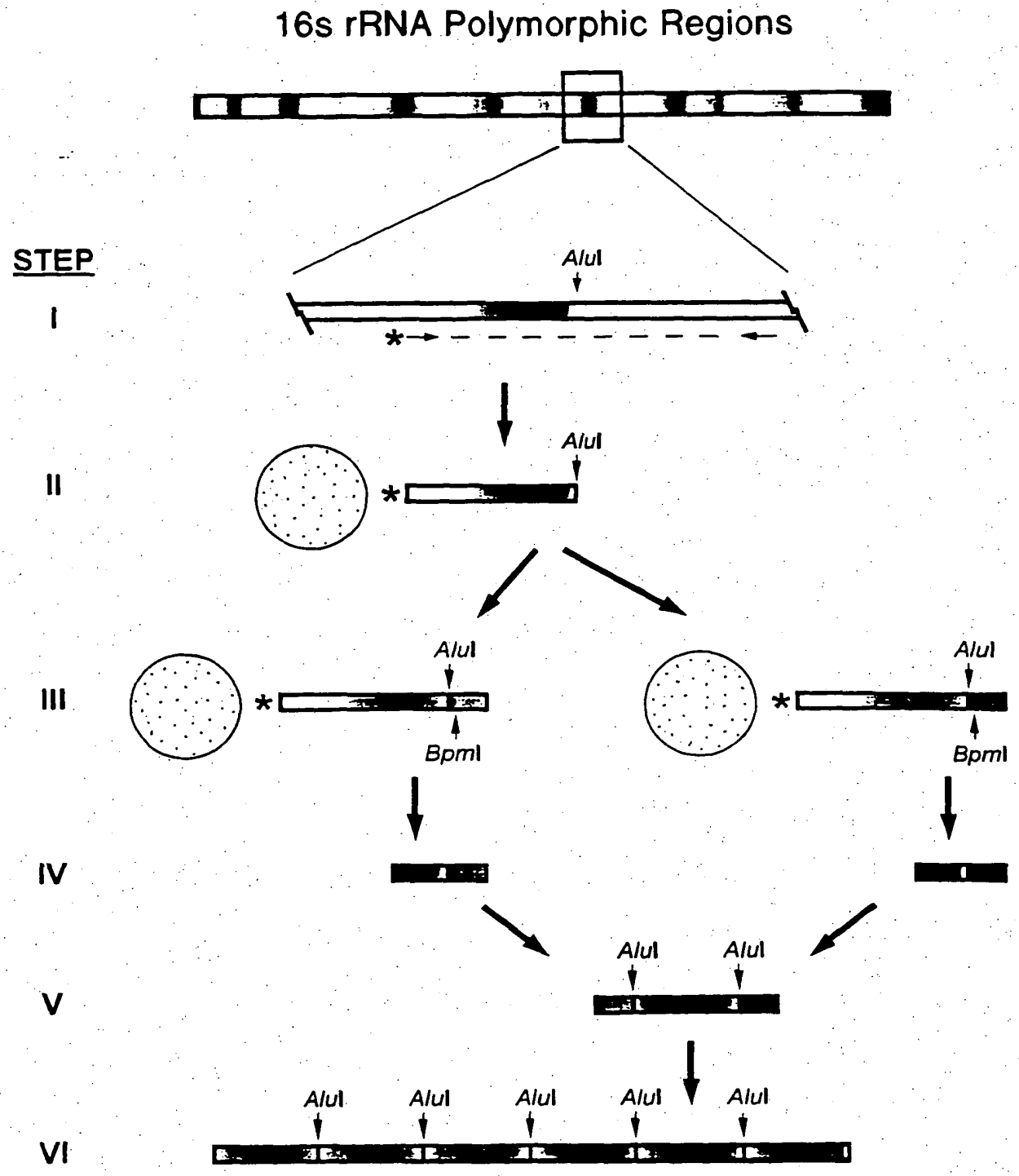


FIG. 2

Bacteria

Acidobacteria Group
Various uncultured environmental Acidobacteria

Aquificales
Desulfurobacterium Group
Desulfurobacterium
Desulfurobacterium thermolithotrophum
environmental samples
unidentified Aquificales OPS132

CFB/Green sulfur bacteria group
Bacteroidaceae
Bacteroides
Bacteroides caccae

Firmicutes(gram-positive bacteria)
Actinobacteria(high G+C gram-positive bacteria)
Actinomycetaceae
Actinomyces
Actinomyces bovis
Actinomyces meyeri
Denitrobacterium
Denitrobacterium detoxificans

Green non-sulfur bacteria
environmental samples
uncultured HC-seep bacterium BPC110
uncultured HC-seep bacterium GCA004
uncultured HC-seep bacterium GCA112

Proteobacteria(purple non-sulfur bacteria)
alpha subdivision
Acetobacteraceae
Acetobacter
Acetobacter aceti
Gluconobacter
Gluconobacter asaii

beta subdivision
Burkholderia group
Burkholderia
***Burkholderia* sp. JB1**
Denitrobacter
Denitrobacter permanens

delta subdivision
Desulfobacter
Desulfobacter curvatus
Desulfobulbus
***Desulfobulbus* sp. BG25**
Legionellaceae
Legionella
Legionella anisa

unclassified Bacteria
benzene mineralizing clone SB-1
uncultured eubacterium env.OPS 1

FIG. 3

- **Archaea (archaebacteria)**
  - **Crenarchaeota (extremely thermophilic archaebacteria)**
    - Desulfurococcales
      - Desulfurococcaceae
        - Aeropyrum
          - *Aeropyrum pernix*
        - Desulfurococcus
          - *Desulfurococcus mobilis*
        - Staphylothermus
          - *Staphylothermus marinus*
    - Sulfolobales
      - Metallosphaera
        - *Metallosphaera sedula*
      - Sulfolobus
        - *Sulfolobus acidocaldarius*
        - *Sulfolobus metallicus*
    - Thermoproteales
      - Thermoproteaceae
        - Caldivirga
          - *Caldivirga maquilingensis*
        - Pyrobaculum
          - *Pyrobaculum islandicum*
  - **Euryarchaeota**
    - Archaeoglobales
      - Archaeoglobaceae
        - Archaeoglobus
          - *Archaeoglobus fulgidus*
          - *Archaeoglobus veneficus*
    - Halobacteriales
      - Halobacteriaceae
        - Haloarcula
          - *Haloarcula japonica*
        - Halococcus
          - *Halococcus morrhuae*
    - Methanococcales
      - Methanococcaceae
        - Methanococcus
          - *Methanococcus jannaschii*
    - Methanobacteriales
      - Methanobacteriaceae
        - Methanobacterium
          - *Methanobacterium bryantii*
          - *Methanobacterium subterraneum*
    - Thermococcales
      - Thermococcaceae
        - Pyrococcus
          - *Pyrococcus abyssi*
    - Thermoplasmales
      - Picrophilaceae
        - Picrophilus
          - *Picrophilus oshimae*

FIG. 4

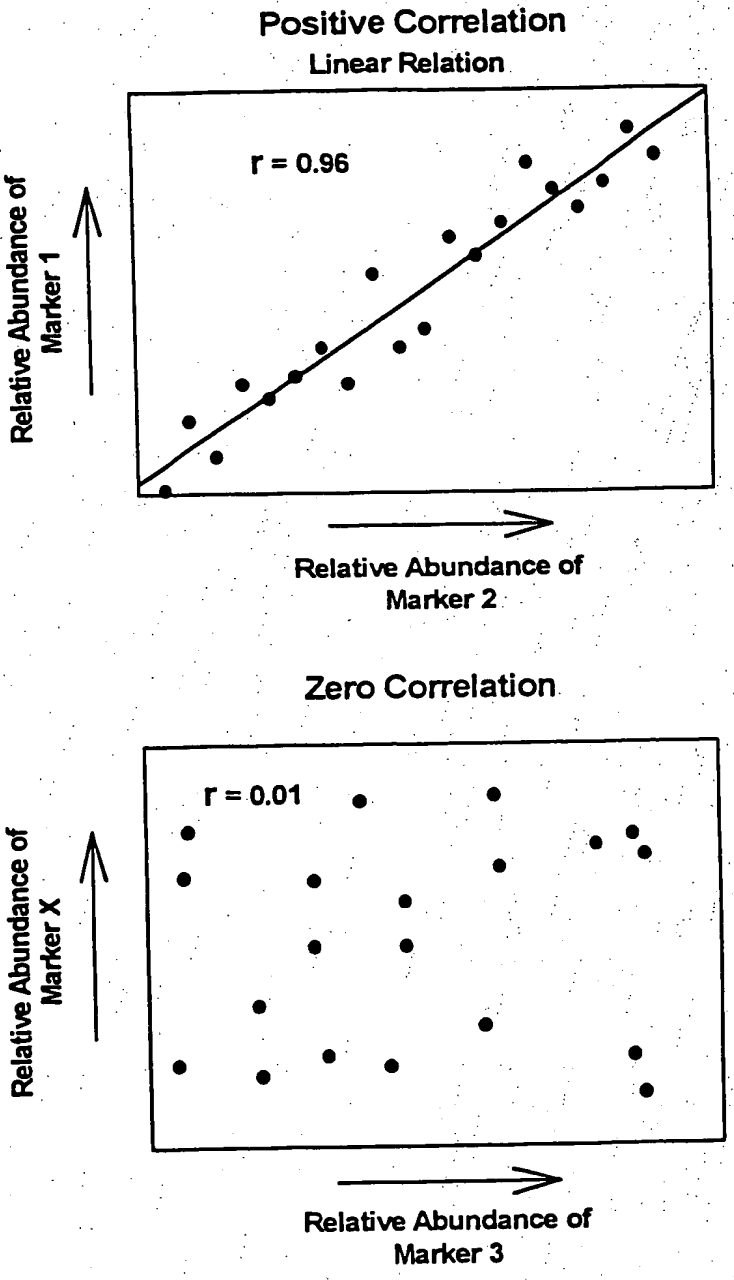

Positive Correlation
Linear Relation
r = 0.96
Relative Abundance of Marker 1
Relative Abundance of Marker 2
Zero Correlation
r = 0.01
Relative Abundance of Marker X
Relative Abundance of Marker 3

FIG. 5

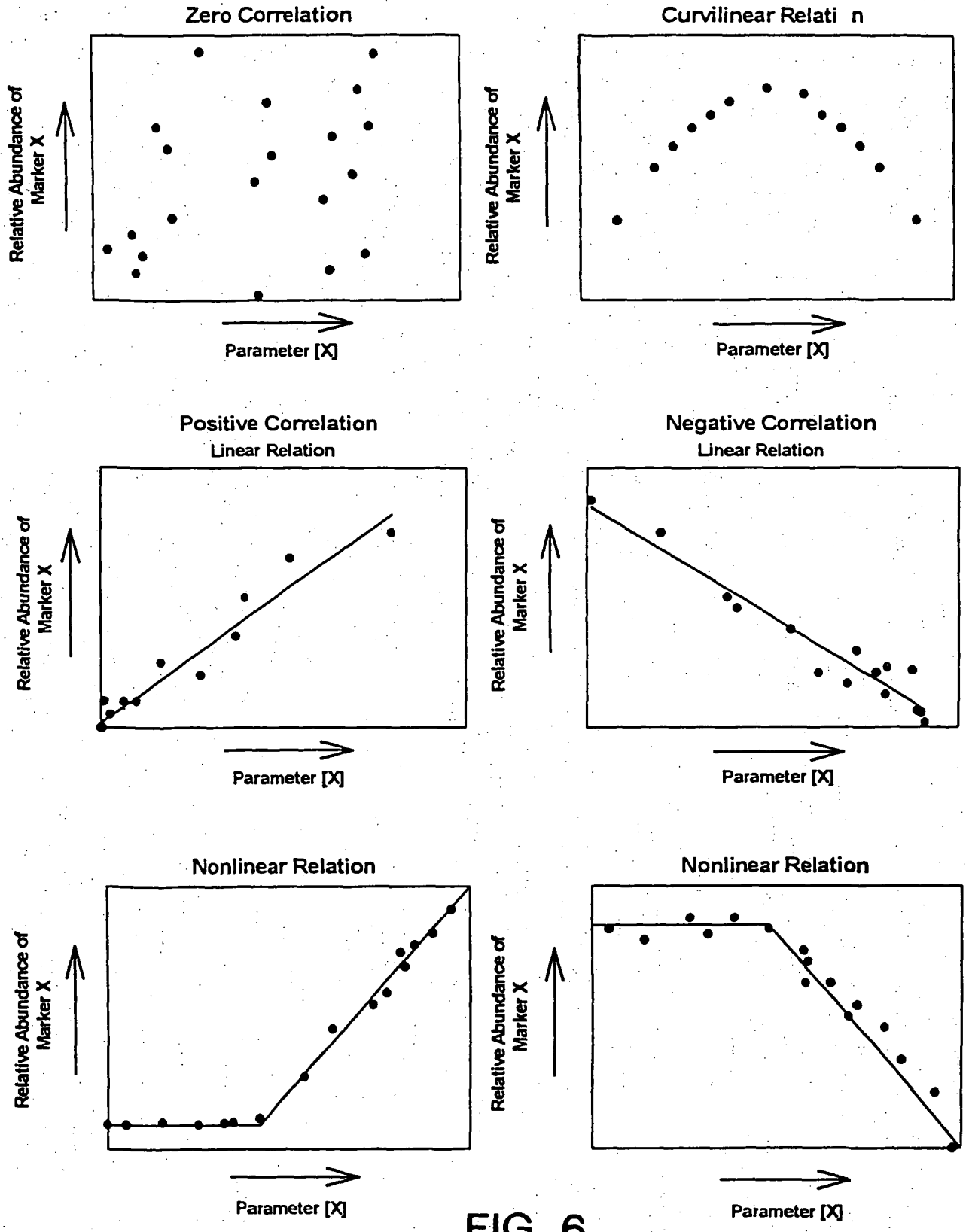

Zero Correlation
Relative Abundance of Marker X
Parameter [X]
Curvilinear Relati n
Relative Abundance of Marker X
Parameter [X]
Positive Correlation
Linear Relation
Relative Abundance of Marker X
Parameter [X]
Negative Correlation
Linear Relation
Relative Abundance of Marker X
Parameter [X]
Nonlinear Relation
Relative Abundance of Marker X
Parameter [X]
Nonlinear Relation
Relative Abundance of Marker X
Parameter [X]

FIG. 6

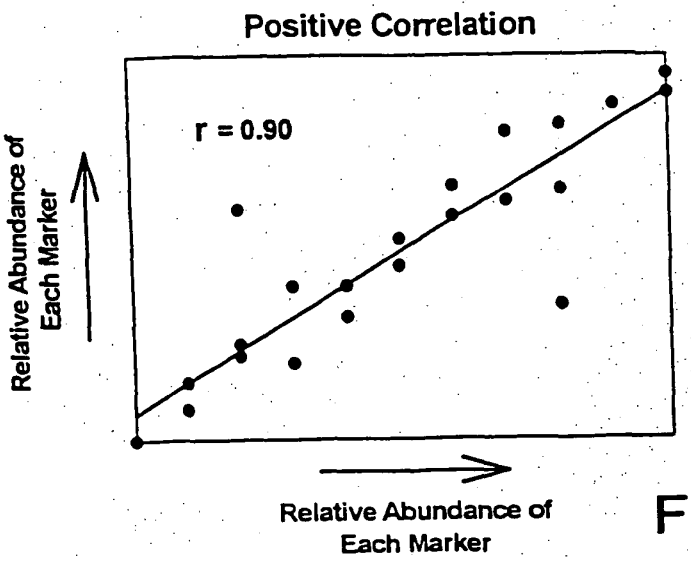


FIG. 7A

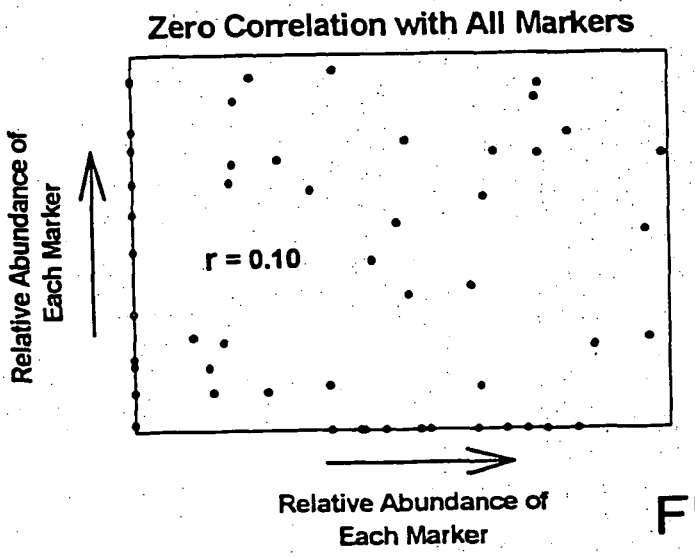


FIG. 7B

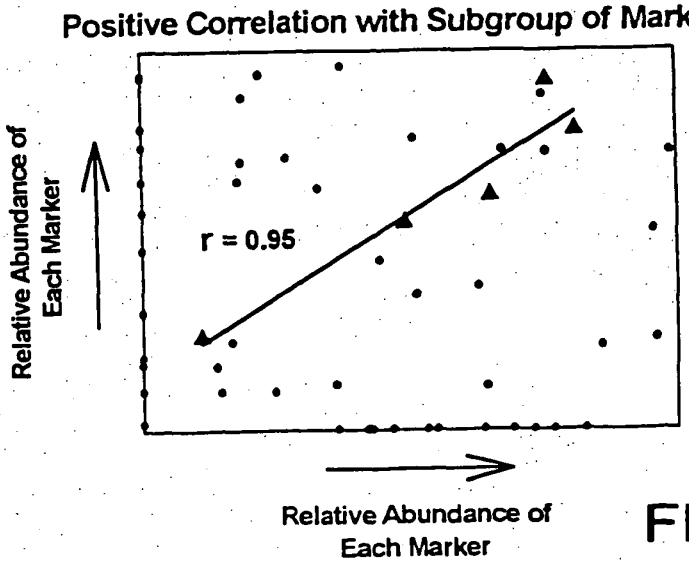


FIG. 7C

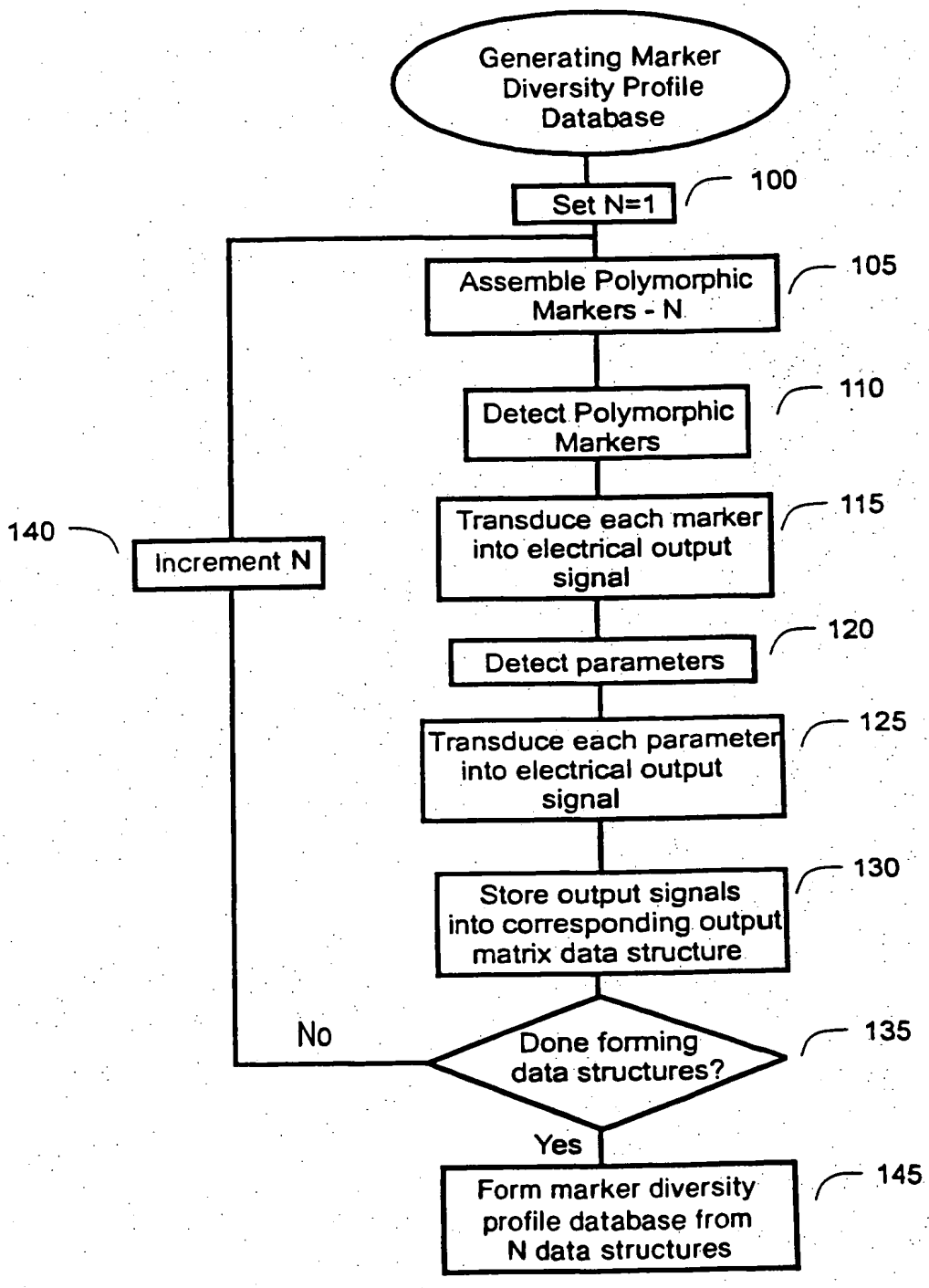

Generating Marker Diversity Profile Database
Set N=1
100
Assemble Polymorphic Markers - N
105
Detect Polymorphic Markers
110
Transduce each marker into electrical output signal
115
Detect parameters
120
Transduce each parameter into electrical output signal
125
Store output signals into corresponding output matrix data structure
130
Done forming data structures?
135
No
Yes
Increment N
140
Form marker diversity profile database from N data structures
145

FIG. 8

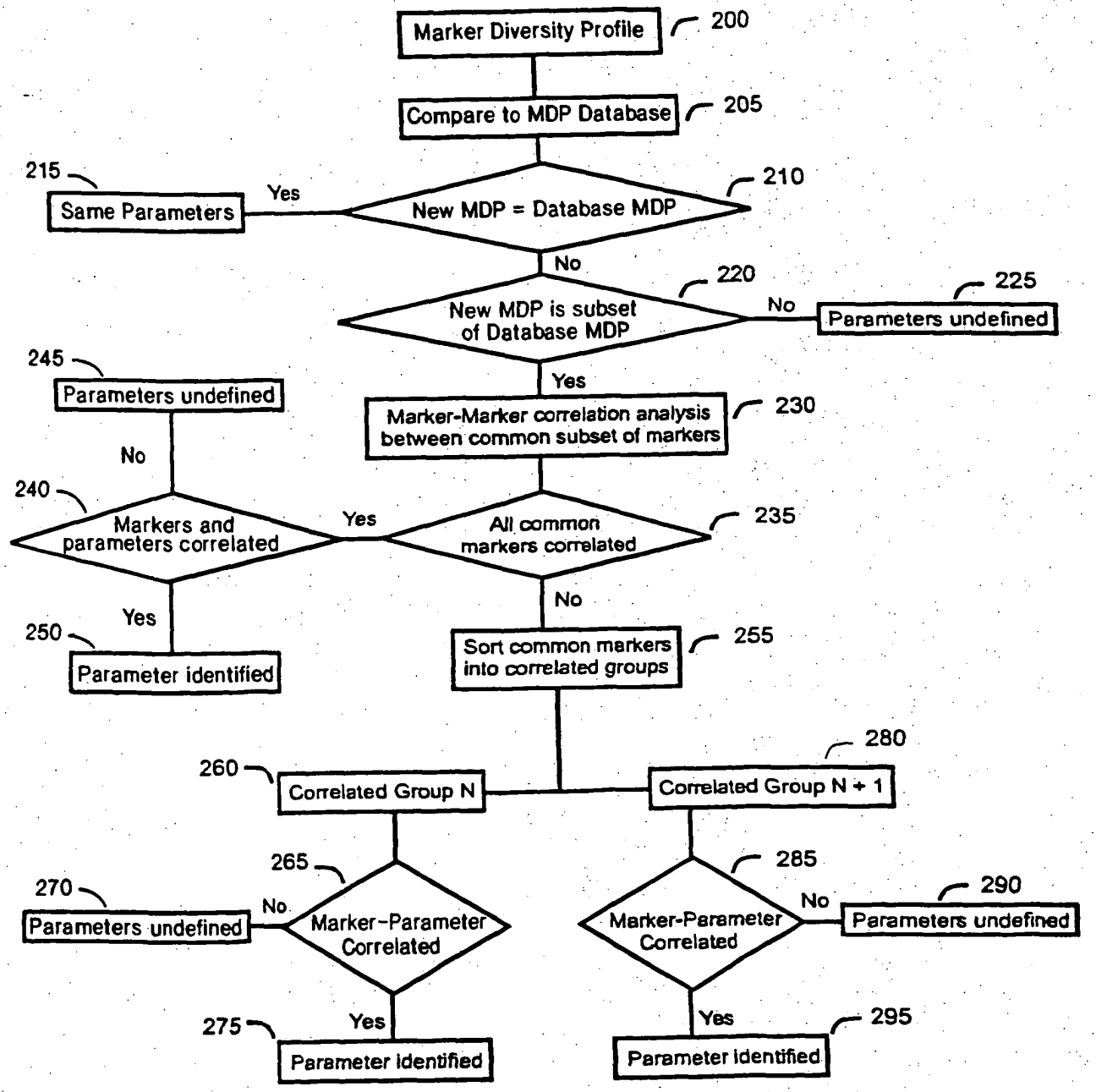

Marker Diversity Profile
200
Compare to MDP Database
205
215
Same Parameters
Yes
New MDP = Database MDP
210
No
220
New MDP is subset of Database MDP
No
225
Parameters undefined
Yes
245
Parameters undefined
Marker-Marker correlation analysis between common subset of markers
230
No
240
Markers and parameters correlated
Yes
All common markers correlated
235
Yes
250
Parameter identified
No
Sort common markers into correlated groups
255
280
260
Correlated Group N
Correlated Group N + 1
265
285
270
Parameters undefined
No
Marker-Parameter Correlated
Marker-Parameter Correlated
No
290
Parameters undefined
275
Yes
Parameter identified
Yes
295
Parameter identified

FIG. 9

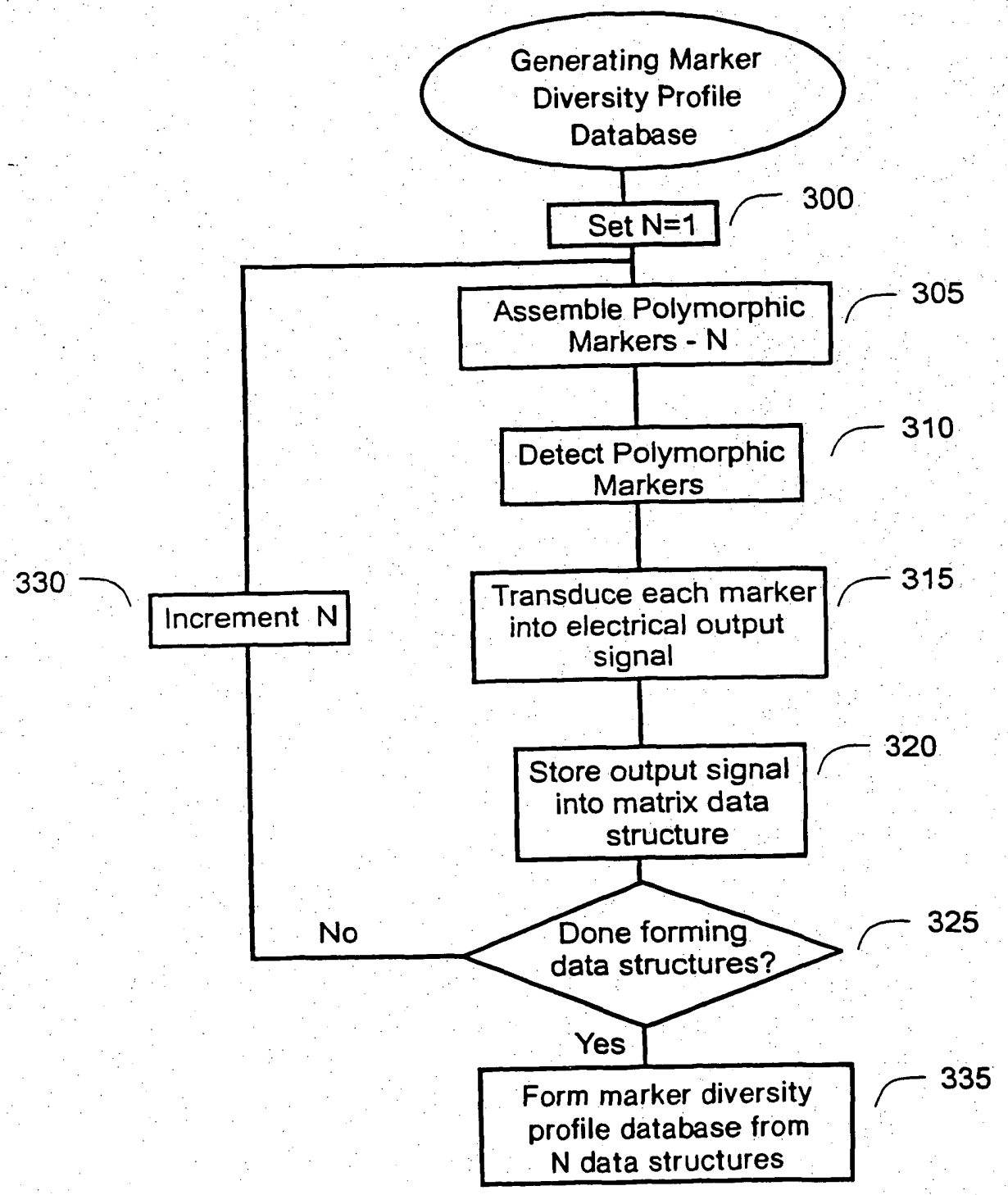

Generating Marker Diversity Profile Database
Set N=1
300
Assemble Polymorphic Markers - N
305
Detect Polymorphic Markers
310
Transduce each marker into electrical output signal
315
Store output signal into matrix data structure
320
Done forming data structures?
325
No
Yes
Increment N
330
Form marker diversity profile database from N data structures
335

FIG. 10

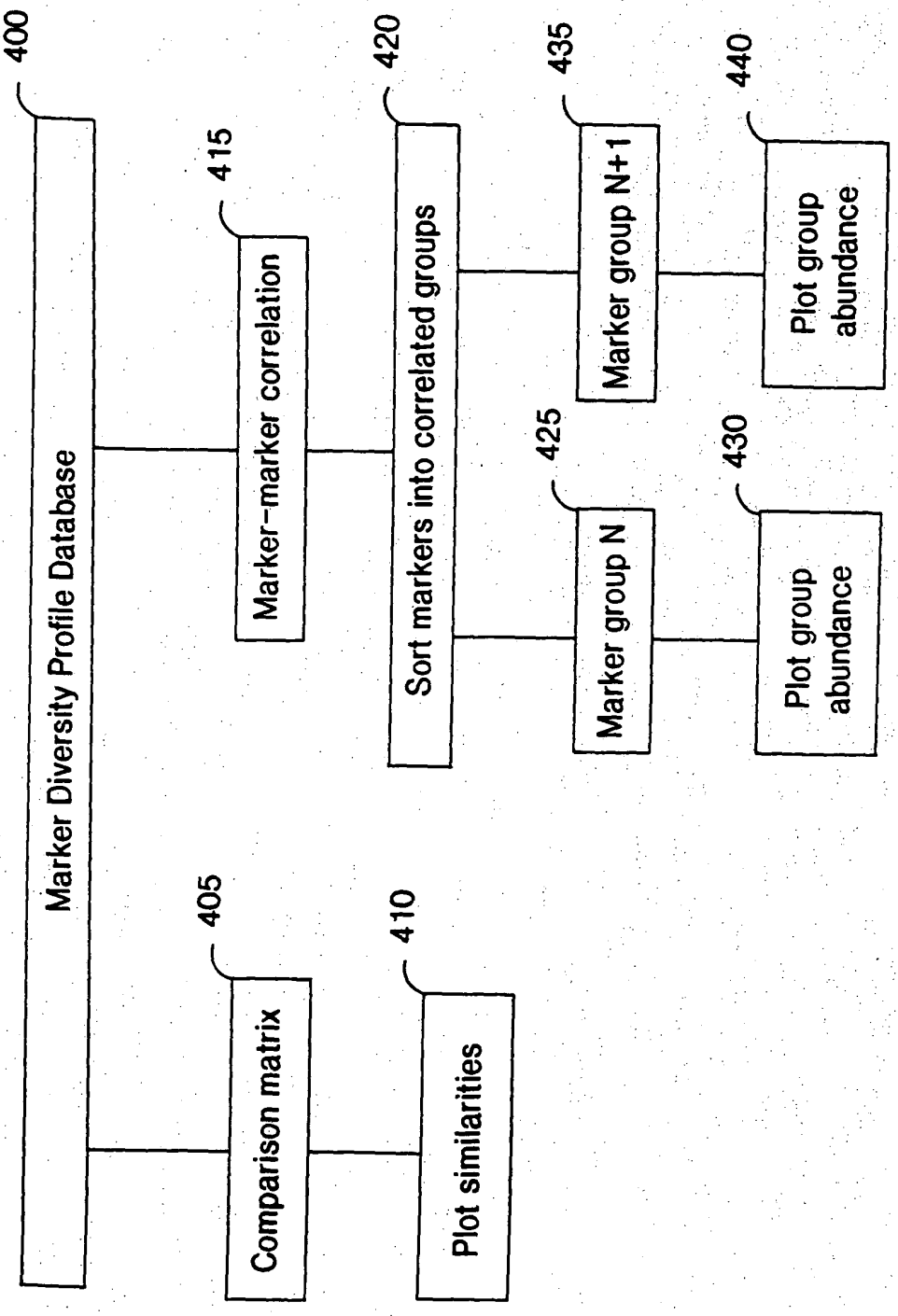

400
Marker Diversity Profile Database
405
Comparison matrix
410
Plot similarities
415
Marker–marker correlation
420
Sort markers into correlated groups
425
Marker group N
430
Plot group abundance
435
Marker group N+1
440
Plot group abundance

FIG. 11

SARD OLIGONUCLEOTIDES

```
TX-003   5'-  Phosphate-GCTCCAGGTCTACATCCTAGTCAGGACddC-3'
TX-012   5'-  Phosphate- CTCCAGGTCTACATCCTAGTCAGGACddC
TX-010R  3'              GAGGTCCAGATGTAGGATCAGTCCTGGATA-5'
TX-013   3'              GAGGTCCAGATGTAGGATCAGTCCTGGATA-5'
TX-004R  3'                    CAGATGTAGGATCAGTCCTGGATA-5'
TX-007R  3'-                   CAGATGTAGGATCAGTCCTGG -Biotin-5'
TX-111   3'-                   CAGATGTAGGATCAGTCCTGGCACTGATAGCTC -Biotin-5'

TX-005   5'-  Phosphate-GCTCCAGACTAGCATCCGCTGACTTGAddC
TX-014   5'-  Phosphate- CTCCAGACTAGCATCCGCTGACTTGAddC
TX-011R  3'              GAGGTCTGATCGTAGGCGACTGAACTGTAA
TX-015   3'              GAGGTCTGATCGTAGGCGACTGAACTGTAA
TX-006R  3'                    TGATCGTAGGCGACTGAACTGTAA
TX-008R  3'                    TGATCGTAGGCGACTGAACTG -Biotin-5'
TX-121   3'                    TGATCGTAGGCGACTGAACTGGCAACGTTGGAC -Biotin-5'


TX-001  5'- Biotin-GTG TAG HRG TGA AAT DCD YA (SEQ ID NO:138)

TX-002  5'- YTC ACG RCA YGA GCT GAC GAC (SEQ ID NO:139)

TX-003  5'- Phosphate-GCT CCA GGT CTA CAT CCT AGT CAG GACddC (SEQ ID NO:140)

TX-004  5'- ATA GGT CCT GAC TAG GAT GTA GAC (SEQ ID NO:141)

TX-005  5'- Phosphate-GCT CCA GAC TAG CAT CCG CTG ACT TGAddC (SEQ ID NO:142)

TX-006  5'- AAT GTC AAG TCA GCG GAT GCT AGT (SEQ ID NO:143)

TX-007  5'- Biotin-GGT CCT GAC TAG GAT GTA GAC (SEQ ID NO:144)

TX-008  5'- Biotin-GTC AAG TCA GCG GAT GCT AGT (SEQ ID NO:145)

TX-009  5'- Biotin-GGA TTA GAW ACC CBG GTA GTC (SEQ ID NO:146)

TX-010  5'- ATA GGT CCT GAC TAG GAT GTA GAC CTG GAG (SEQ ID NO:147)

TX-011  5'  AAT GTC AAG TCA GCG GAT GCT AGT CTG GAG (SEQ ID NO:148)

TX-012  5'- Phosphate-CTC CAG GTC TAC ATC CTA GTC AGG AcddC (SEQ ID NO:149)

TX-013  5'- ATA GGT CCT GAC TAG GAT GTA GAC CTG GAG (SEQ ID NO:150)

TX-014  5'- Phosphate-CTC CAG ACT AGC ATC CGC TGA CTT GaddC (SEQ ID NO:151)

TX-015  5'  AAT GTC AAG TCA GCG GAT GCT AGT CTG GAG (SEQ ID NO:152)

TX-111  5'  Biotin-CTC GAT AGT CAC GGT CCT GAC TAG GAT GTA GAC (SEQ ID NO:153)

TX-121  5'  Biotin-CAG GTT GCA ACG GTC AAG TCA GCG GAT GCT AGT (SEQ ID NO:154)
```

FIG. 12

SARD STRATEGY

1. Amplify rDNA with TX-009 and 1392R

```
Bio---------nnnnNNNNNNNNNNNNAGCTnnnnnnnnnnnnnnnnn
   ---------nnnnNNNNNNNNNNNNTCGAnnnnnnnnnnnnnnnnn
```

2. Cut with *AluI*.

```
Bio---------nnnnNNNNNNNNNNNNAG    CTnnnnnnnnnnnnnnnn
   ---------nnnnNNNNNNNNNNNNTC    GAnnnnnnnnnnnnnnnn
```

3. Split into two pools, bind to SA beads, wash, add TX-12/13 adapter *or* TX-14/15 adapter.

```
Bio---------nnnnNNNNNNNNNNNNAG    CTCCAGGTCTACATCCTAGTCAGGACddC -3' (TX012)
   ---------nnnnNNNNNNNNNNNNTC    GAGGTCCAGATGTAGGATCAGTCCTGGATA-5' (TX013)

                                OR

Bio---------nnnnNNNNNNNNNNNNAG    CTCCAGACTAGCATCCGCTGACTTGAddC -3' (TX014)
   ---------nnnnNNNNNNNNNNNNTC    GAGGTCTGATCGTAGGCGACTGAACTGTAA-5' (TX015)
```

4. Ligate adapters, wash.

```
Bio---------nnnnNNNNNNNNNNNNAGCTCCAGGTCTACATCCTAGTCAGGACddC -3' (TX012)
   ---------nnnnNNNNNNNNNNNNTCGAGGTCCAGATGTAGGATCAGTCCTGGATA-5' (TX013)
```

FIG. 13

5. Cleave with *Bpm*I, remove SA-beads.

```
     NNNNNNNNNNNNAGCTCCAGGTCTACATCCTAGTCAGGACddC -3' (TX012)
nnNNNNNNNNNNNTCGAGGTCCAGATGTAGGATCAGTCCTGGATA-5' (TX013)
```

6. Remove 3' termini with PolI plus 4 dNTPs, combine pools, ligate.

```
(TX015)                                                                    (TX012)
AATGTCAAGTCAGCGGATGCTAGTCTGGAGCTNNNNNNNNNNNNNNNNNNNNNNNNAGCTCCAGGTCTACATCCTAGTCAGGACddC -3'
 ddCAGTTCAGTCGCCTACGATCAGACCTCGANNNNNNNNNNNNNNNNNNNNNNNNTCGAGGTCCAGATGTAGGATCAGTCCTGGATA-5'
(TX014)                                                                    (TX013)
```

7. Amplify ditags with TX-111, TX-121 primers

```
BioCAGGTTGCAACGGTCAAGTCAGCGGATGCTAGT---------->
           AATGTCAAGTCAGCGGATGCTAGTCTGGAGCTNNNNNNNNNNNNNNNNNNNNNNNNAGCTCCAGGTCTACATCCTAGTCAGGACddC -3
            ddCAGTTCAGTCGCCTACGATCAGACCTCGANNNNNNNNNNNNNNNNNNNNNNNNTCGAGGTCCAGATGTAGGATCAGTCCTGGATA-5'
                                                  <------CAGATGTAGGATCAGTCCTGGCACTGATAGCTC-Bio
```

8. Purify PCR product, cleave with *Alu*I

```
BioCAGGTTGCAACGGTCAAGTCAGCGGATGCTAGTCTGGAG  CTNNNNNNNNNNNNNNNNNNNNNNAG  CTCCAGGTCTACATCCTAGTCAGGACCGTGACTATCGAG-3'
   GTCCAACGTTGCCAGTTCAGTCGCCTACGATCAGACCTC  GANNNNNNNNNNNNNNNNNNNNNNTC  GAGGTCCAGATGTAGGATCAGTCCTGGCACTGATAGCTC-Bio
```

9. Purify away adapters, ligate.

```
CTNNNNNNNNNNNNNNNNNNNNNNNNAGCTNNNNNNNNNNNNNNNNNNNNNNNNAGCTNNNNNNNNNNNNNNNNNNNNNNNNAG
GANNNNNNNNNNNNNNNNNNNNNNNNTCGANNNNNNNNNNNNNNNNNNNNNNNNTCGANNNNNNNNNNNNNNNNNNNNNNNNTC
```

10. G I-purify large concatemers, clone into *Sma*I site of vector.

FIG. 13 (cont.)

Applicants: Matthew Ashby
Application No.: Not yet assigned
For: METHODS FOR THE SURVEY AND GENETIC ANALYSIS OF POPULATIONS
Agent Lawrence M. Brown, Reg. No. 52,660
Docket No.: ASHBY/1 DIV
Filed: June 25, 2003
EV133107841 US

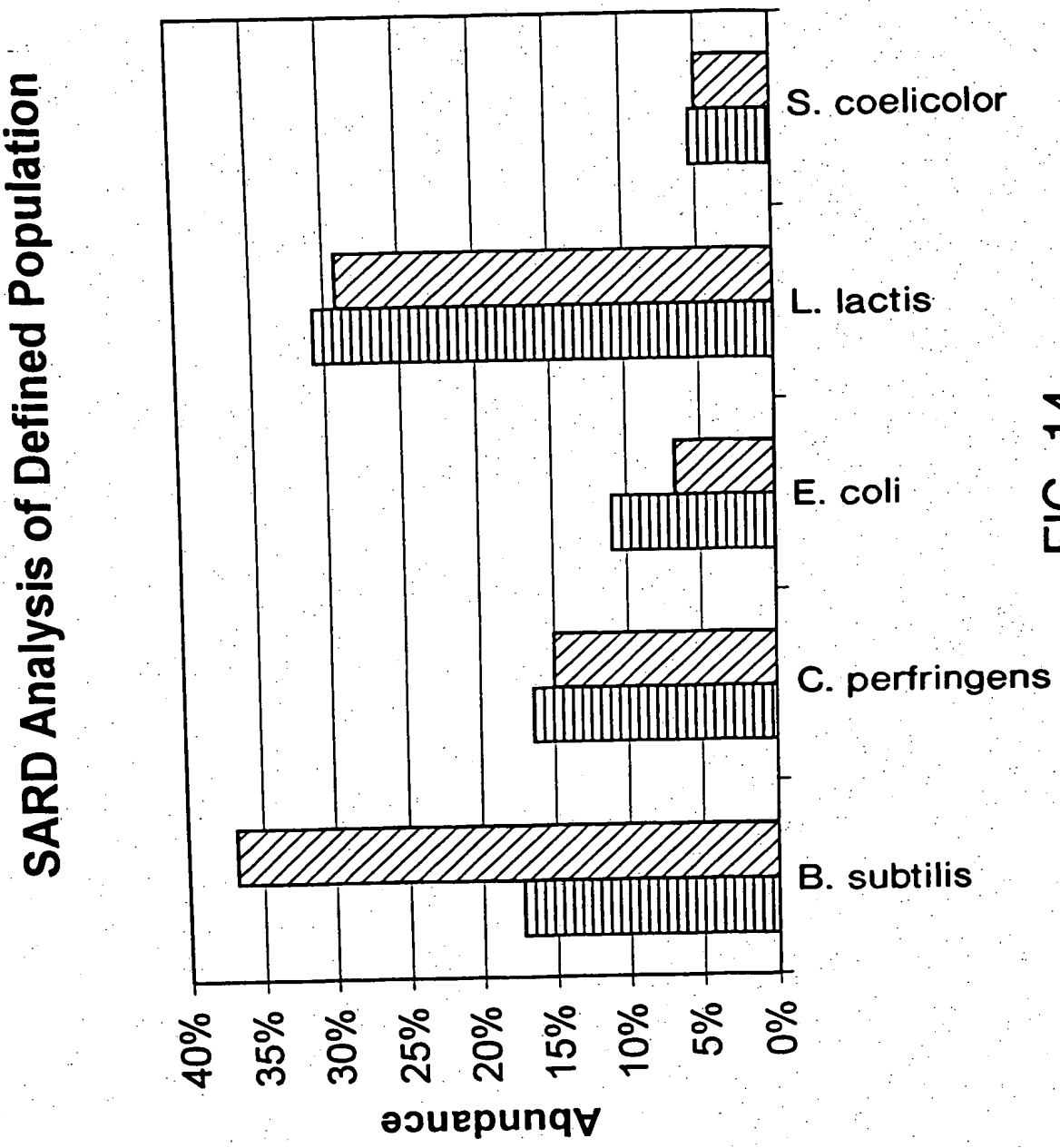

SARD Analysis of Defined Population
S. coelicolor
L. lactis
E. coli
C. perfringens
B. subtilis
40%
35%
30%
25%
20%
15%
10%
5%
0%
Abundance

FIG. 14

ACGATGAGCACTAGCT (SEQ ID NO:1) (1)

ACGATGAGTACTAGCT (SEQ ID NO:2) (1)

ACGATGATGACTAGCT (SEQ ID NO:3) (2)

ACGATGGATGCTAGCT (SEQ ID NO:4) (1)

ATGCTAGTCTGGAGCT (SEQ ID NO:5) (6)

ATGGCTGTCGTCAGCT (SEQ ID NO:6) (50)

ATGGTTGTCGTCAGCT (SEQ ID NO:7) (1)

ATTCCGTGCCGTAGCT (SEQ ID NO:8) (1)

CACTAGTGGCGCAGCT (SEQ ID NO:9) (2)

CCCCCGTGCCGAAGCT (SEQ ID NO:10) (1)

CCCCCGTGCCGCAGCT (SEQ ID NO:11) (1)

CCCCCTTCCTCCAGCT (SEQ ID NO:12) (1)

CCCCGGTGCCGCAGCT (SEQ ID NO:13) (1)

CCGGGTAGTCCCAGCT (SEQ ID NO:14) (5)

CCTCCGTGCCGAAGCT (SEQ ID NO:15) (3)

CCTCCGTGCCGCAGCT (SEQ ID NO:16) (2)

CCTCCGTGCTGCAGCT (SEQ ID NO:17) (1)

CCTCGGCGCCGCAGCT (SEQ ID NO:18) (2)

CCTCGGTGCCGCAGCT (SEQ ID NO:19) (1)

CCTCGGTGTCGCAGCT (SEQ ID NO:20) (2)

CCTGGGTGCCGCAGCT (SEQ ID NO:21) (2)

CCTGTGTGACGAAGCT (SEQ ID NO:22) (1)

CCTTGGTAACGAAGCT (SEQ ID NO:23) (1)

CCTTGGTACCGAAGCT (SEQ ID NO:24) (1)

CGCCAGTGCCGTAGCT (SEQ ID NO:25) (1)

CGCCTGTGCCGTAGCT (SEQ ID NO:26) (2)

CGTCCGTGCCGAAGCT (SEQ ID NO:27) (2)

CGTCCGTGCCGCAGCT (SEQ ID NO:28) (1)

CGTCGGTGCCGCAGCT (SEQ ID NO:29) (3)

CTCCCGTGCCGCAGCT (SEQ ID NO:30) (1)

CTCCCGTGCCGGAGCT (SEQ ID NO:31) (1)

CTCCGGTGCCGCAGCT (SEQ ID NO:32) (1)

CTCCTGTGCCGAAGCT (SEQ ID NO:33) (1)

CTCCTGTGCCGCAGCT (SEQ ID NO:34) (2)

CTGCCGTGCCGAAGCT (SEQ ID NO:35) (9)

CTGCTGTGCCGAAGCT (SEQ ID NO:36) (2)

CTGTCGTGCCGAAGCT (SEQ ID NO:37) (1)

CTTCAGTATCGAAGCT (SEQ ID NO:38) (1)

CTTCCGCGCCGGAGCT (SEQ ID NO:39) (2)

CTTCCGTGCCGCAGCT (SEQ ID NO:40) (3)

CTTCCGTGCCGGAGCT (SEQ ID NO:41) (1)

CTTCGGTGCCGCAGCT (SEQ ID NO:42) (1)

CTTCTGTGGCGAAGCT (SEQ ID NO:43) (1)

GATCCGTGCCGTAGCT (SEQ ID NO:44) (1)

GCTCTGTGCCGAAGCT (SEQ ID NO:45) (1)

GCTGGGTGCCCAAGCT (SEQ ID NO:46) (1)

GGTCCGTGCCGCAGCT (SEQ ID NO:47) (1)

GGTGCTCTTCGGAGCT (SEQ ID NO:48) (2)

GTAAACGATGGAAGCT (SEQ ID NO:49) (1)

GTGGCTGTCGTCAGCT (SEQ ID NO:50) (2)

GTTCCGTGCCGAAGCT (SEQ ID NO:51) (2)

GTTCCGTGCCGCAGCT (SEQ ID NO:52) (3)

TATCAGTGGCGCAGCT (SEQ ID NO:53) (1)

TCTCCGTGCCGCAGCT (SEQ ID NO:54) (1)

TCTCTGTGCCGCAGCT (SEQ ID NO:55) (2)

TCTCTGTGCCGTAGCT (SEQ ID NO:56) (1)

TGTCCGTGCCGTAGCT (SEQ ID NO:57) (1)

TTTCCGTGCCGCAGCT (SEQ ID NO:58) (2)

FIG. 15

ACGATGATAACTAGCT (SEQ ID NO:59) (2)

ACGATGATGACTAGCT (SEQ ID NO:60) (1)

ACGATGGGCACTAGCT (SEQ ID NO:61) (1)

ACGGCTGTCGTCAGCT (SEQ ID NO:62) (2)

ACTACGAGCGCAAGCT (SEQ ID NO:63) (1)

ACTTAATGCGTTAGCT (SEQ ID NO:64) (1)

ATGCTAGTCTGGAGCT (SEQ ID NO:65) (8)

ATGGCTCTCGTCAGCT (SEQ ID NO:66) (1)

ATGGCTGTCGCCAGCT (SEQ ID NO:67) (3)

ATGGCTGTCGTCAGCT (SEQ ID NO:68) (70)

ATGGTTGTCGTCAGCT (SEQ ID NO:69) (2)

ATGTAGACCTGGAGCT (SEQ ID NO:70) (12)

ATTCCGTGCCGCAGCT (SEQ ID NO:71) (6)

ATTCCGTGCCGTAGCT (SEQ ID NO:72) (1)

CACAAGCGGTGGAGCT (SEQ ID NO:73) (1)

CACTAGTGGCGCAGCT (SEQ ID NO:74) (4)

CATCCGTGCCGAAGCT (SEQ ID NO:75) (1)

CCCCAGGGCCCAAGCT (SEQ ID NO:76) (1)

CCCCGGTGCCGCAGCT (SEQ ID NO:77) (1)

CCCGCGTGCCGGAGCT (SEQ ID NO:78) (1)

CCGCGGTGCCGTAGCT (SEQ ID NO:79) (1)

CCGGGTAGTCCCAGCT (SEQ ID NO:80) (4)

CCGGGTAGTCCTAGCT (SEQ ID NO:81) (1)

CCTCCGTGCCGAAGCT (SEQ ID NO:82) (2)

CCTCCGTGCCGCAGCT (SEQ ID NO:83) (2)

CCTCCGTGCCGGAGCT (SEQ ID NO:84) (1)

CCTCGTAAGGGGAGCT (SEQ ID NO:85) (1)

CCTGGGTGCCGCAGCT (SEQ ID NO:86) (1)

CCTGGTAGTCCCAGCT (SEQ ID NO:87) (2)

CCTGGTAGTCCTAGCT (SEQ ID NO:88) (1)

CCTTAGTAACGCAGCT (SEQ ID NO:89) (1)

CCTTGGTAACGAAGCT (SEQ ID NO:90) (1)

CGCCAGTGCCGAAGCT (SEQ ID NO:91) (1)

CGCCGGTGCCGCAGCT (SEQ ID NO:92) (2)

CGCCTGTGCCGTAGCT (SEQ ID NO:93) (2)

CGCTCGTGGCGAAGCT (SEQ ID NO:94) (1)

CGGAGGCGTCGTAGCT (SEQ ID NO:95) (1)

CGTCAGTGTCGCAGCT (SEQ ID NO:96) (2)

CGTCCGTGCCGAAGCT (SEQ ID NO:97) (7)

CGTCCGTGCCGCAGCT (SEQ ID NO:98) (1)

CGTCCGTGCCGGAGCT (SEQ ID NO:99) (1)

CGTCGGTGCCGCAGCT (SEQ ID NO:100) (2)

CTCCAGTGCCGCAGCT (SEQ ID NO:101) (1)

CTCCCGTGCCACAGCT (SEQ ID NO:102) (1)

CTCCCGTGCCGCAGCT (SEQ ID NO:103) (2)

CTCCCGTGCCGGAGCT (SEQ ID NO:104) (3)

CTCCGGTGCCGCAGCT (SEQ ID NO:105) (1)

CTCCTGTGCCGCAGCT (SEQ ID NO:106) (1)

CTGCCGCGCCGGAGCT (SEQ ID NO:107) (1)

CTGCCGTGCCGAAGCT (SEQ ID NO:108) (5)

CTGCCGTGCCTAAGCT (SEQ ID NO:109) (1)

CTGCGGTGCCGCAGCT (SEQ ID NO:110) (1)

CTGCTGTGCCGAAGCT (SEQ ID NO:111) (7)

CTGTCGTGCCGAAGCT (SEQ ID NO:112) (1)

CTTCCGCGCCGGAGCT (SEQ ID NO:113) (7)

CTTCCGTGCCGAAGCT (SEQ ID NO:114) (2)

CTTCCGTGCCGCAGCT (SEQ ID NO:115) (6)

CTTCCGTGCCGGAGCT (SEQ ID NO:116) (2)

CTTCGGTGCCGCAGCT (SEQ ID NO:117) (1)

CTTCGGTGTCGCAGCT (SEQ ID NO:118) (1)

CTTGGGTGCCGCAGCT (SEQ ID NO:119) (2)

CTTTAGTAACGCAGCT (SEQ ID NO:120) (2)

GACCCGCAAGGGAGCT (SEQ ID NO:121) (1)

GATCCGTGCCGCAGCT (SEQ ID NO:122) (2)

GCTCCGTGCCGAAGCT (SEQ ID NO:123) (1)

GCTCCGTGCCGTAGCT (SEQ ID NO:124) (1)

GCTCTGTGCCGAAGCT (SEQ ID NO:125) (2)

GCTCTGTGCCGTAGCT (SEQ ID NO:126) (1)

GGGCTTGTCGTCAGCT (SEQ ID NO:127) (1)

GTAAACGATGGAAGCT (SEQ ID NO:128) (1)

GTTCCGTGCCGCAGCT (SEQ ID NO:129) (2)

GTTCCGTGCCGTAGCT (SEQ ID NO:130) (2)

GTTCTGTGCCGCAGCT (SEQ ID NO:131) (2)

TCTCACGACACGAGCT (SEQ ID NO:132) (2)

TCTCAGTAACGTAGCT (SEQ ID NO:133) (1)

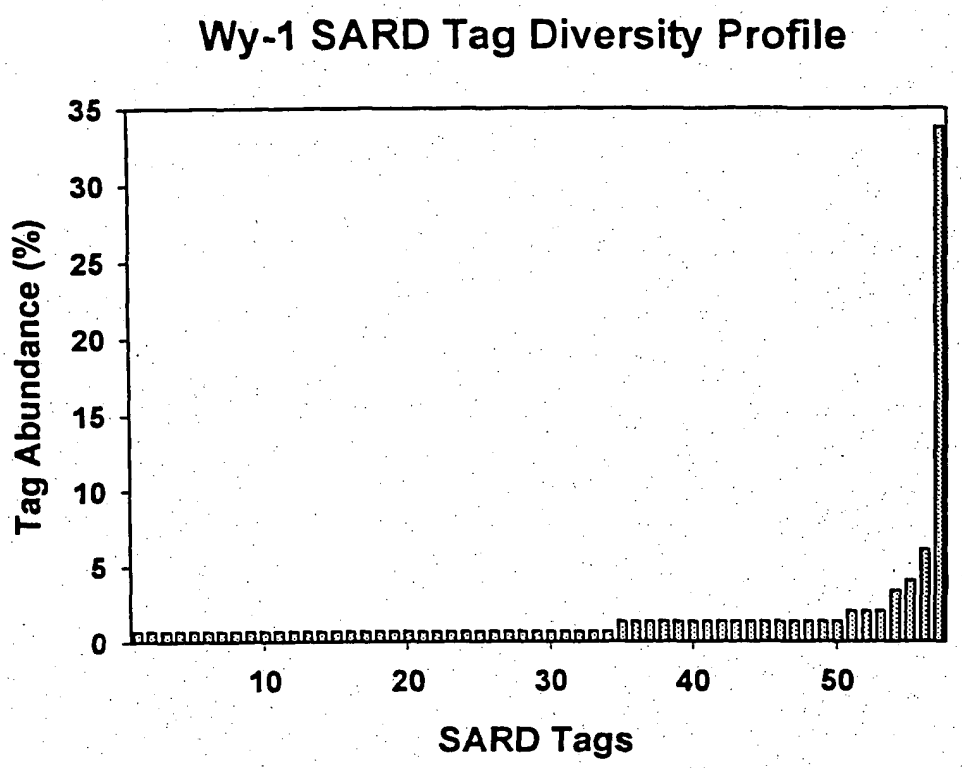
TCTCCGTGCCGCAGCT (SEQ ID NO:134) (2)

TCTCTGTGCCGCAGCT (SEQ ID NO:135) (4)

TGGACGTTGCGGAGCT (SEQ ID NO:136) (2)

TTTCCGTGCCGGAGCT (SEQ ID NO:137) (1)

FIG. 16


Wy-1 SARD Tag Diversity Profile
Tag Abundance (%)
0
5
10
15
20
25
30
35
10
20
30
40
50
SARD Tags

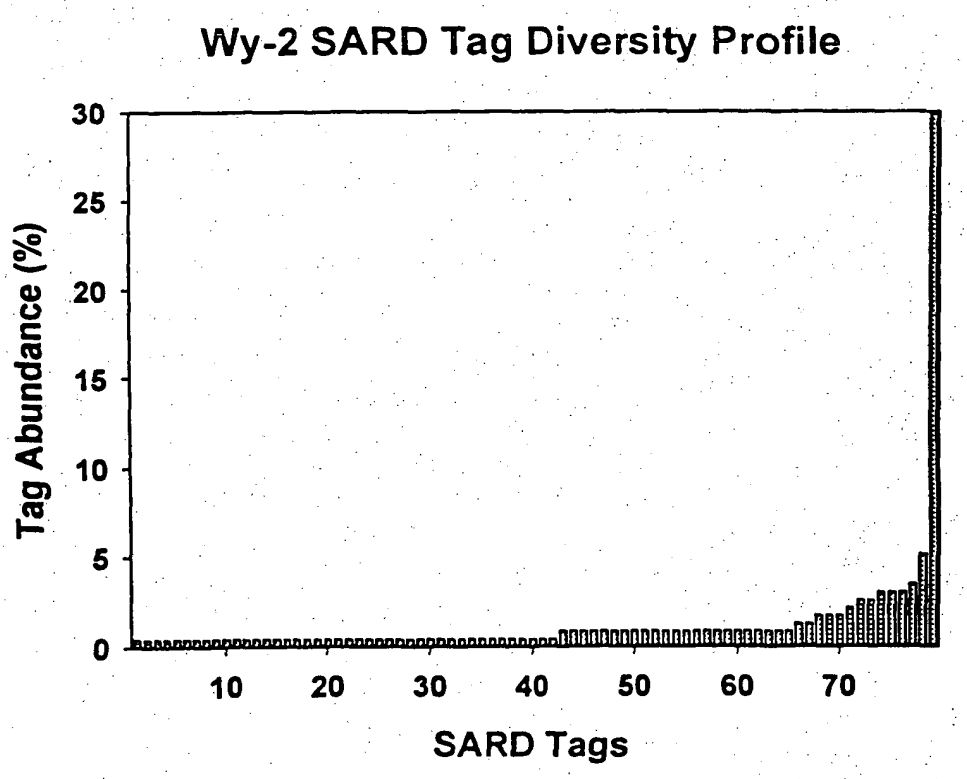

Wy-2 SARD Tag Diversity Profile
Tag Abundance (%)
0
5
10
15
20
25
30
10
20
30
40
50
60
70
SARD Tags

FIG. 17